Catalogue no. 75F0002MIE — No. 007
ISSN: 1707-2840
ISBN: 0-662-43805-1

Research Paper

Income Research Paper Series

Housing Data in the Survey of Labour and Income Dynamics (SLID)

by Heather Lathe

Income Statistics Division
Jean Talon Building, Ottawa, ON K1A 0T6

Telephone: 1 613 951-7355

Statistics Canada Statistique Canada

Canada

How to obtain more information

Specific inquiries about this product and related statistics or services should be directed to: Income Statistics Division, Statistics Canada, Ottawa, Ontario, K1A 0T6 (telephone: (613) 951-7355; (888) 297-7355; income@statcan.ca).

For information on the wide range of data available from Statistics Canada, you can contact us by calling one of our toll-free numbers. You can also contact us by e-mail or by visiting our website.

National inquiries line	1 800 263-1136
National telecommunications device for the hearing impaired	1 800 363-7629
Depository Services Program inquiries	1 800 700-1033
Fax line for Depository Services Program	1 800 889-9734
E-mail inquiries	infostats@statcan.ca
Website	www.statcan.ca

Information to access the product

This product, catalogue no. 75F0002MIE, is available for free. To obtain a single issue, visit our website at www.statcan.ca and select Our Products and Services.

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner and in the official language of their choice. To this end, the Agency has developed standards of service that its employees observe in serving its clients. To obtain a copy of these service standards, please contact Statistics Canada toll free at 1 800 263-1136. The service standards are also published on www.statcan.ca under About Statistics Canada > Providing services to Canadians.

Statistics Canada
Income Statistics Division

Income Research Paper Series

Housing Data in the Survey of Labour and Income Dynamics (SLID)

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2006

All rights reserved. The content of this electronic publication may be reproduced, in whole or in part, and by any means, without further permission from Statistics Canada, subject to the following conditions: that it be done solely for the purposes of private study, research, criticism, review or newspaper summary, and/or for non-commercial purposes; and that Statistics Canada be fully acknowledged as follows: Source (or "Adapted from", if appropriate): Statistics Canada, year of publication, name of product, catalogue number, volume and issue numbers, reference period and page(s). Otherwise, no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form, by any means—electronic, mechanical or photocopy—or for any purposes without prior written permission of Licensing Services, Client Services Division, Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

July 2006

Catalogue no. 75F0002MIE, Vol. 7

Frequency: Occasional

ISSN: 1707-2840
ISBN: 0-662-43805-1

Ottawa

La version française de cette publication est disponible sur demande (nº 75F0002MIF au catalogue).

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Abstract

This paper summarizes data available from the Survey of Labour and Income Dynamics (SLID) on housing characteristics and shelter costs, with a special focus on imputation methods used. From 1994 to 2001, the survey covered only a few housing characteristics, primarily ownership status and dwelling type. In 2002, thanks to the sponsorship of Canada Mortgage and Housing Corporation (CMHC), several other characteristics and detailed shelter costs were added to the survey. Several imputation methods were also introduced at that time, in order to replace missing values due to survey non-response and to provide utility costs, which contribute to total shelter costs. These methods take advantage of SLID's longitudinal design and also use data from other sources such as the Labour Force Survey and the Census. In July 2006, further improvements in the imputation methods were introduced for 2004 and applied to past years in a historical revision. This report also documents that revision.

Acknowledgements

The author wishes to thank Chris Duddek and Susan Storey in Household Survey Methods Division and Steve Grunau, Willa Rea, Preston Poon and Anne-Marie Vinette in Income Statistics Division for their contributions and comments.

Table of Contents

1. Introduction 7
2. Participation of Canada Mortgage and Housing Corporation 8
3. Summary of available data 9
4. Missing values replacement: Longitudinal imputation 11
5. Missing values replacement: Cross-sectional or "donor" imputation 13
6. Revision of housing data from 1999 to 2004 14
7. Imputation rates and revision results 15
8. Imputation of utility costs 19
9. Where to find more information 21
References 22

1. Introduction

The Survey of Labour and Income Dynamics (SLID) is an household survey conducted by Statistics Canada via telephone interviews each year. It is designed to measure the economic well-being of Canadians, including changes they experience over time and the factors affecting these changes.

Housing characteristics and shelter costs of Canadians have become a main component of SLID quite recently. Data on household (or family) incomes, demographic characteristics, labour market participation, and a variety of other content areas have been part of the survey since it began in 1993. Most of the housing content was developed starting with the 2002 reference year (collected in January 2003) at the request of Canada Mortgage and Housing Corporation (CMHC).[1] CMHC required annual income and housing data in a relatively large-sample survey, to supplement data from the Census of Population (every five years) and the Survey of Household Spending. The ability to do longitudinal analysis with SLID data was also of interest to CMHC. At the time of this report, the annual sponsorship by CMHC of housing content in SLID is continuing every year.

This report summarizes the data available from SLID on housing characteristics and shelter costs, with a special focus on the imputation methods used for this data. Imputation plays an important role in most household surveys as a means of replacing what would otherwise be missing values in the output data. Missing values arise from non-response or errors in the data collected from respondents. Although complete non-response to the questionnaire is taken care of through the re-distribution of survey weights, non-response to certain questions can be dealt with during survey processing by imputing (assigning) appropriate valid values in place of "don't know" responses.

1. Canada Mortgage and Housing Corporation (CMHC) is Canada's national housing agency. CMHC works to enhance Canada's housing finance options, assist Canadians who cannot afford housing in the private market, improve building standards and housing construction, and provide policymakers with the information and analysis they need to sustain a vibrant housing market in Canada.

Imputation is used in SLID to replace missing values for all its income and housing content. In addition, SLID uses imputation to cover expenditures on utilities like electricity, home heating and water, which are difficult to collect by telephone interview with a high degree of accuracy. This instance of imputation is called mass imputation because it is used to obtain utility costs for 100% of respondents in the sample.

The rest of this document begins with a brief summary of the role of CMHC in the survey. Then the content is briefly listed, with explanations of how to access additional information from other survey publications on the Statistics Canada web site. Sections 4 and 5 describe in non-technical terms how imputation is currently used to replace missing values in SLID's housing content. A more technical explanation is available in an earlier paper, entitled *General Housing Imputation (excluding utilities) in the Survey of Labour and Income Dynamics* (but some of the tables are out of date). Sections 6 and 7 describe how imputation was re-applied for years 2002 to 2004 shortly after the release of data for 2004 in order to make some improvements, and how the same methods were applied back to 1999 for the variables existing prior to 2002. Section 8 provides a short description of how imputation is used to impute utility costs for all households in order to include such costs in the total shelter cost estimates, since accurate utility costs are difficult to collect by telephone interview.

2. Participation of Canada Mortgage and Housing Corporation

Canada Mortgage and Housing Corporation (CMHC) assesses housing conditions and develops estimates of housing need using data extracted every five years from the Census of Population. To monitor housing conditions between censuses, CMHC uses survey data, for example, information from the Survey of Household Spending (SHS), and before 1997, from the Household Facilities and Equipment Survey (HFE). This allows them to have annual trends and a greater choice of other characteristics with which to cross housing data on Canadian households.

In 2001, CMHC began to sponsor additional content in both the SLID and SHS, starting with reference year 2002. The additional content was particularly significant for SLID. Until then, SLID had only covered a few housing characteristics, primarily ownership status and dwelling type. Furthermore, no imputation of those housing characteristics was done. Over 20 housing-related questions were added to the SLID interview. Some questions apply to homeowners and renters, some only to homeowners, and some only to renters. Only one person answers these questions on behalf of all members of the household.

3. Summary of available data

This section contains a list of the housing characteristics available for analysis. Associated with each content item, or "variable", is the short variable name used for programming purposes, a short label to better identify what it means, and the question (or questions) by which it is collected in the questionnaire. Longer variable descriptions and category descriptions can be found on the internet in the SLID data dictionary. Precise question wording and the order that questions are asked in the interview can be seen by referring to the research paper that contains the Entry-Exit module of the questionnaire. See Section 9, "Where to find more information".

As explained above, the majority of housing content in SLID started in 2002, thanks to sponsorship of CMHC.

The first reference year of SLID, 1993, contained no variables on housing. For reference years 1994 to 1998, two questions were added, corresponding to dwelling tenure (DWLTEN25) and the 4-category dwelling type (DWLTYP25). Starting with 1999, another two new questions began to be collected, partly as a result of computer-assisted questionnaire standardization between surveys. The new questions covered the number of bedrooms in the dwelling and, if the owner was a member of the household, whether or not they had a mortgage.

Also starting with reference year 1999, the response categories for the question on dwelling type was expanded from four categories (single detached, semi-detached, townhouse, and apartment) to seven, as in the variable DWLDET25. For many analytical purposes it is sufficient to have just the four broad categories of dwelling type, and therefore DWLTYP25 is a collapsed version of DWLDET25 as of 1999.

Variable name	Label	Question number
Available from 1999:		
DWLDET25	Detailed dwelling type (9 categories)	DWT_Q01; CON-Q6C before 1999
DWLTYP25	Dwelling type (4 categories)	Derived; CON-Q6D before 1999
DWTENR25	Dwelling tenure, i.e., ownership status	TN_Q01
ROOMS25	Number of bedrooms	BD_Q01
MORTG25	Mortgage on property (yes/no)	MTGE_Q01
Available from 2002:		
REPA25	Dwelling repairs needed	OR_Q005
HEAT25*	Principal heating fuel (also, HEATG25)	OR_Q010
OPFM25	Household operates farm from property	OR_Q015
OPBU25**	Household operates business from property	OR_Q020
PRTXI25**	Property taxes incl. in mortgage payments	OW_Q005
PRTXM25	Monthly property taxes	OW_Q045
MORTGN25	More than one mortgage on the dwelling	OW_Q010
MORTGP25**	Periodicity of mortgage payments	OW_Q015
MORTGM25	Monthly mortgage payment excl. prop. Tax	OW_Q025 – Q040
COND25	Dwelling is a condominium	OW_Q050
CONDM25	Monthly condominium fee	OW_Q055
RENTM25	Monthly rent (excl. utilities)	RN_Q005
RNPK25	Parking included in rent	RN_Q010
RNHT25	Heating fuel included in rent	RN_Q010
RNWA25	Water included in rent	RN_Q010
RNEC25	Electricity included in rent	RN_Q010
RNTV25*	Cable TV included in rent	RN_Q010
RNFG25*	Refrigerator included in rent	RN_Q010
RNST25*	Cooking stove included in rent	RN_Q010
RNWD25*	Washer and dryer included in rent	RN_Q010
RNFU25	Furniture included in rent	RN_Q010
RNNO25	No amenities included in rent	RN_Q010
FURN25*	Rent includes partially or fully furnished	RN_Q010
RNRE25	Reduced rent (yes/no and reason)	RN_Q020
RNBS25	Rent calc. on the basis of income (yes/no)	RN_Q025
UHTY25, UWAY25, UECY25	Utility costs for heating fuel, water and electricity, respectively	Imputed, all values
SHELY25	Total annual shelter costs	Derived by summing
RQNU25, SUIT25	Required number of bedrooms for the members, and whether this number is sufficient	Derived***

* Available until 2005 only

** Missing values due to non-response are left as such in the output data for years 2002-2005, but will be replaced with imputed values starting in 2006.

*** From Canada Mortgage and Housing Corporation, "Special Studies on 1996 Census Data: Canadian Housing Conditions." Research Highlights: Socioeconomic Series, Issue 55-1, p. 1, 4: A dwelling is suitable if the number of bedrooms in the dwelling is equal or greater than the required number of bedrooms as specified by the National Occupancy Standard.

Starting with reference year 2006, two new questions have been added. Both are important in the imputation of utility costs (described in Section 8), which contribute to total shelter costs. A new question addressed to owners asks whether payments for electricity, heating fuel, or water are included in other payments such as condominium fees. If the other payments already include this expense, we can impute $0 for those utilities. A new question addressed to renters who report $0 for monthly rent payments asks whether they pay for any utilities and, if so, which ones. If they do pay for utilities, we will impute a positive expenditure amount.

Also starting with reference year 2006, the dwelling type question is broken down into about four short questions to gather the same information in a more telephone-friendly way (SLID is a telephone survey while the LFS uses a mixture of telephone and "personal", i.e., face-to-face, interviews). The previous long list of possible categories, which had to be read out loud to the respondent over the telephone, was not received favourably.[2]

4. Missing values replacement: Longitudinal imputation

Longitudinal imputation is used in surveys that have some longitudinal aspect to their collection, in that they interview the same respondents for more than one collection period. It takes into account, as much as possible, information already collected from the survey respondent during a previous contact. SLID follows respondents for up to six years. Its respondents were also participants in the Labour Force Survey (LFS) one year before entering SLID—this provides yet another opportunity to use longitudinal imputation.

The usefulness of longitudinal imputation is greatly increased when the characteristics to be imputed are expected to remain similar from one period to the next based on certain

2. It may also have elicited an overly high proportion of responses to the first category, "single-detached", as compared to the Census and SHS which do not have a telephone collection mode. When the 2006 data for SLID become available, it will be possible to assess whether the modified question brings SLID dwelling type results more in line with the Census and SHS, or whether an underlying problem of non-response bias is the cause.

other characteristics. For example, we can take advantage of the fact that many housing characteristics are expected to be the same if the respondent did not move. The postal code can be used to determine, at least with a reasonable probability, that the respondent did not move.

The postal code is obtained as part of the respondent's address each time they are interviewed. If there is a contact with the respondent in a particular year, the postal code was likely obtained, even if the interview was not completed in full. Most respondents who were not contacted have weights re-distributed among the respondent households, so that their final weights are zero, meaning that they are excluded from the estimates and therefore do not need to be imputed.

The LFS and the Rent Schedule, which is a supplement of the LFS, were used as sources of information for longitudinal imputation for respondents in their first year of SLID. Missing SLID information was imputed using the household's LFS information if the household did not move. The four variables that could be imputed this way were dwelling type, dwelling tenure, the number of rooms, and (for owners) whether there was a mortgage.

To use the Rent Schedule, it was necessary to ensure that the dwelling tenure (ownership) variable on SLID indicated that the household was still a renter. The variables whose values are potentially taken from the Rent Survey in longitudinal imputation include the indicators of what amenities (utilities, parking, etc.) are included in the household's rental payments, and the amount of monthly rent.

In addition to bringing forward LFS or Rent Schedule data, SLID data from the previous year are also used to impute data for subsequent years, if the household does not move. Imputed data can potentially get carried forward for more than one year, but households who are completely non-respondent to the survey for successive years generally have their weight set to zero for those years.

5. Missing values replacement: Cross-sectional or "donor" imputation

Briefly stated, donor imputation consists of identifying a group of households sharing several characteristics with the household whose data must be imputed and copying the data from one household into the missing field of the household to be imputed. The shared input characteristics are selected from the other variables in the survey. For example, SLID imputes the income of an individual by picking the income of a person living in the same province and with the same age, sex, level of education, type of job (employee, self-employed) and occupation.

In the case of SLID's housing variables, the variables to be imputed are treated in separate steps or modules, depending on whether they apply to all households or only certain types of households such as those who rent their home, those who own, owners with a mortgage (mortgage-related variables), etc. Not all households require complete imputation of housing characteristics, in which case they may be dealt with in just one or two of the imputation modules. The modules are structured so that even if a household does require complete imputation of housing characteristics, the variables imputed in earlier modules can then be used as inputs in others.

The first module serves to impute dwelling type and dwelling tenure. If a household needs imputation for either of these variables, it probably needs complete imputation because these are among the first questions asked about housing in the SLID questionnaire. The input variables for grouping donor households and households to be imputed are the province, urban size ranges (not described here) and the number of persons in the household. This is just the first step in the donor imputation of housing characteristics in SLID. More information is provided in the technical research paper, *General Housing Imputation (excluding utilities) in the Survey of Labour and Income Dynamics*.

6. Revision of housing data from 1999 to 2004

The publication of this report was timed to accompany a re-issue of the SLID data on housing for the years 1999 to 2004. The remaining SLID content is the same as released on March 30, 2006.

One of the reasons for the housing revision was the re-imputation of utility costs. The second reason was to apply the missing values imputation to more years: the actual techniques have not changed, but they have been expanded back to 1999 and re-applied more fully for 2002-2004, in the interest of increasing the accuracy and usefulness of the collected data. It was decided not to carry the imputation methods back to earlier years for operational reasons.

Both longitudinal and donor imputation have now been carried out on the data for 1999 to 2001. Specifically, longitudinal imputation was used for respondents entering the survey in 1999 (Panel 3 respondents) by taking their LFS data for 1998, including the LFS Rent Schedule. Longitudinal imputation using SLID previous-year data was then applied to all three years (Panel 2 respondents only for 1999). Some backwards longitudinal imputation, i.e., using subsequent-year data, was also possible, due to the fact that this was a historical revision. Finally, donor imputation was applied to all three years to replace the remaining missing values[3].

The replacement of missing values for 1999-2001 made it possible to apply, for the first time, longitudinal imputation for Panel 3 respondents in 2002, for the key variables of dwelling type, tenure, number of rooms and presence of a mortgage. While this was being carried out, the opportunity was also taken to improve the longitudinal imputation

3. As a historical note, longitudinal imputation had already been used for the variables on housing for 1999 and 2000 since SLID's 2001 release. This was done to compensate for errors in collection in those years that had to do with the computerized SLID questionnaire. If SLID were not a longitudinal survey, the missing data for dwelling type in 1999 and all four housing characteristics in 2000 would have affected close to 40% of households. Instead, longitudinal imputation brought the non-response rates down to much more acceptable levels. Donor imputation was not applied at that time.

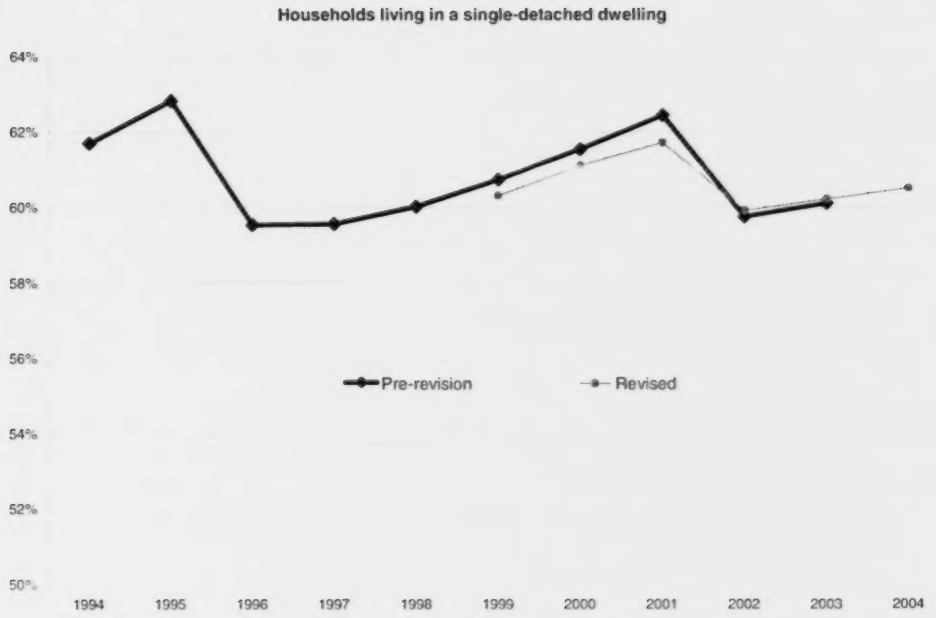
of the Panel 4 respondents in 2002. Specifically, an error in how their 2001 Rent Schedule data was originally accessed caused it to be under-utilized, but this has now been corrected. Longitudinal imputation was then re-applied for 2003 and 2004. The resulting increase in longitudinally imputed information for 2002-2004 then lowered the reliance on donor imputation when it was re-applied for all three years.

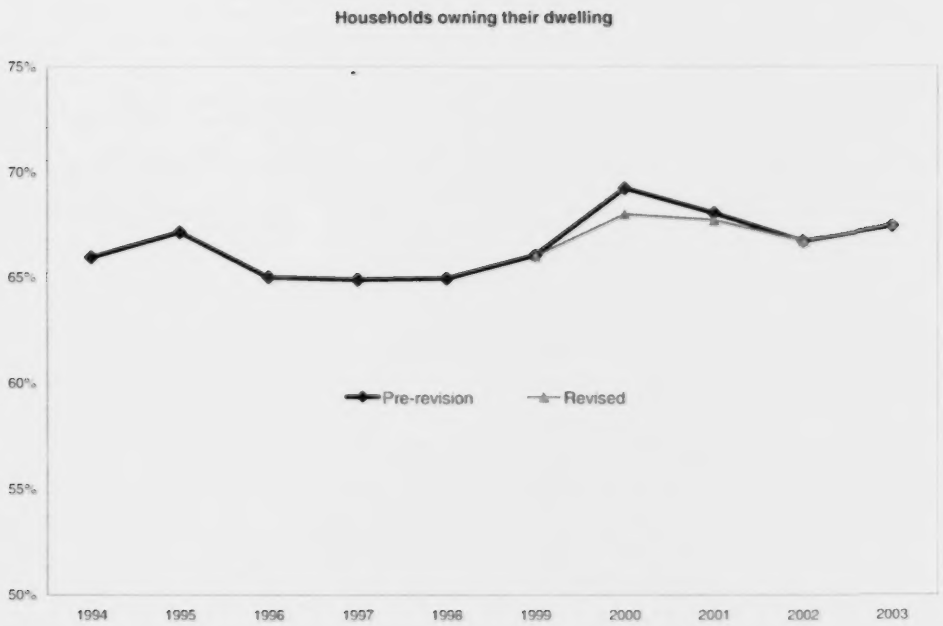
While introducing these improved methods for replacing missing housing values, two additional changes were made. First, the variable indicating whether a farm is operated on the property (OPFM25) was edited for 2003 to make it consistent with other years in terms of the population eligible for a valid value. Second, an edit in 1999 where households with zero bedrooms (ROOMS25) were changed to one bedroom has been reversed.

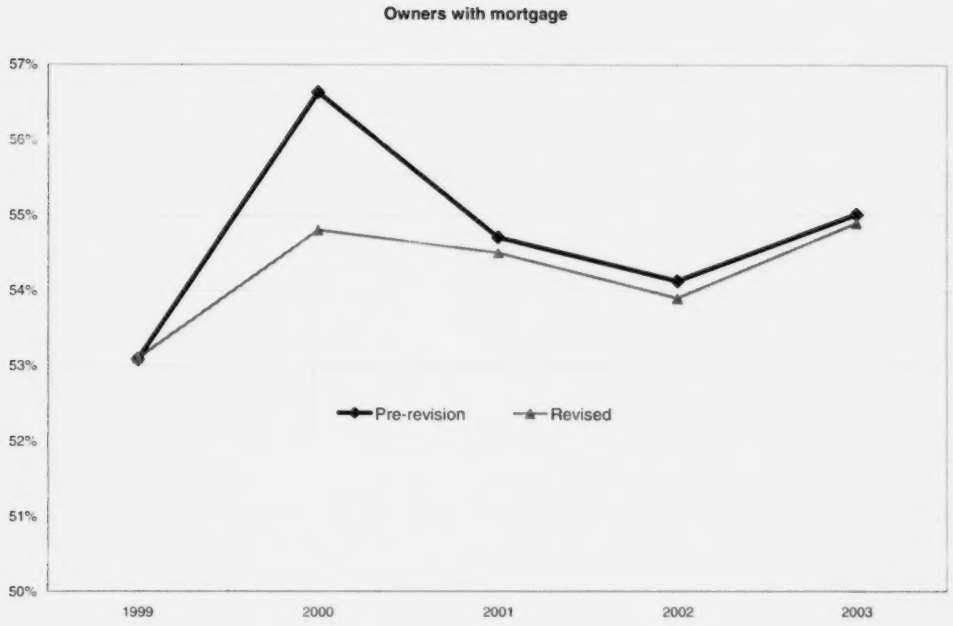
7. Imputation rates and revision results

Following the revision from 1999 to 2004, there are no longer any missing values for dwelling type, tenure, number of rooms and whether the household has a mortgage. Having complete data without missing values increases the data's usefulness at lower levels of aggregation, such as for smaller geographic areas, because there are more records in the sample available for estimates.

The following charts show the pre- and post-revision trends in the percentage of households living in a single-detached dwelling, owning their dwelling, and having a mortgage (owners only), respectively. (The years 1994 to 1998 are included in order to show the longer trend from the survey, although the percentages for 1994-1998 were obtained by simply excluding records with missing values.)


Households living in a single-detached dwelling
64%
62%
60%
58%
56%
54%
52%
50%
Pre-revision
Revised
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004


Households owning their dwelling
75%
70%
65%
60%
55%
50%
Pre-revision
Revised
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003



Based on these three variables, the revision had a favourable impact on the data at aggregate levels, particularly for the four years from 1999 to 2002. The new trends for these three characteristics have been smoothed out somewhat.

However, there still remains some evidence of non-response bias, particularly between 2001 and 2002 which was when approximately half of the sample respondents were replaced as Panel 4 was introduced and Panel 2 was retired. The most likely explanation is the "panel effect" caused by attrition, i.e., the high non-response in the last year of a panel is replaced by much better response as a new panel is introduced. Although imputation for missing values can reduce the impact of non-response bias, it does not necessarily eliminate it.[4]

4. It was hoped that non-response bias would be further minimized by pairing the longitudinal and cross-sectional imputation methods: while the first method can deal with households who did not move, the latter can deal with those that did. The result was not as good as was anticipated, but there are two possible explanations. First, there are relatively few characteristics in the donor imputation modules; at present, only household-level characteristics are used, but some member-level characteristics might be considered in future. Second, the available sample of potential donor records has already been eroded by non-response to the survey, thereby contributing to non-response bias within the donor pool; it might be of interest to take into account final weights at this stage.

Table 1 is an updated version of Table 8.4 provided in *Data Quality in the 2004 Survey of Labour and Income Dynamics (SLID)*, showing partial and total imputation rates. However, unlike the 2004 Data Quality Report, it now excludes households that receive final weights of zero, reflecting the fact that complete non-response to the survey is dealt with through weight adjustments and not imputation. Also, the variable on whether the household operates a farm on the property (OPFM25) has been excluded from the table: the edit to set OPFM25 to "not applicable" for urban households, previously counted as partial imputation, was causing the partial imputation rates to be overstated and the total imputation rates to be understated (except in New Brunswick where it is not applicable for both urban and rural households).

Table 1 - Households requiring imputation of housing variables, 2004

Province	Total imputation	Partial imputation	No imputation
	percent		
Newfoundland	5.2	27.5	67.3
Prince Edward Island	4.5	27.2	68.2
Nova Scotia	5.0	23.0	72.1
New Brunswick	5.6	20.3	74.2
Quebec	6.1	15.4	78.6
Ontario	7.0	24.1	68.9
Manitoba	5.5	26.6	68.0
Saskatchewan	5.0	27.3	67.7
Alberta	6.6	26.7	66.8
British Columbia	7.9	24.7	67.4
Canada	**6.2**	**23.0**	**70.8**

Table 2 shows examples of the extent to which imputation for missing values was necessary and how the imputed values were obtained. For each variable, those households which provided a valid value during the interview did not require imputation. For the remaining households, either longitudinal imputation or donor imputation was used. For 1999 and 2002, longitudinal imputation was possible using the LFS as the data source for the households of Panel 3 and Panel 4, respectively. The rates shown in the

table are based on weighted data, which excludes households that were completely non-respondent to the interview.

Table 2 Distribution of households by source of data, for selected variables and years

	All Households		Renter Households	Households with mortgage
	Dwelling type (DWLTYP25)	**Dwelling tenure** (DWTENR25)	**Monthly rent** (RENTM25)	**Monthly mortgage payment** (MORTGM25)
1999				
As collected	97.10	53.72	...	...
Longitudinal: LFS	0.19	43.48	...	...
Longitudinal: SLID	2.62	0.90	...	...
Donor	0.09	1.90	...	...
2002				
As collected	93.15	93.43	87.91	70.83
Longitudinal: LFS	1.92	1.84	...	...
Longitudinal: SLID	4.51	4.31	2.59	8.82
Donor	0.42	0.42	9.50	20.35
2004				
As collected	92.15	92.30	85.45	71.87
Longitudinal: LFS	...	...	...	...
Longitudinal: SLID	7.45	7.27	10.51	12.85
Donor	0.40	0.43	4.04	15.28

... not applicable

8. Imputation of utility costs

Utility costs (water, fuel and electricity costs) can be a large expense for many households and play a role in affordability issues. However, it is difficult to obtain accurate information from respondents in a telephone interview. Instead of interviewing respondents about these expenditures, SLID uses a method known as mean imputation to add this information, using data from the Census. Since it is used for 100% of the SLID sample, it is also called mass imputation.

The Census collects utility costs on the 20% "long form" sample every five years. This provides a very rich dataset—especially in terms of detailed communities—from which to obtain imputation values. In addition to the detailed breakdown of the sample by community, a large number of other characteristics are used as inputs to impute utility expenditures, namely: dwelling type, dwelling tenure (ownership), whether it is a condominium, the number of bedrooms, the number of persons in the household and household income.

For renter households, it is also possible to find a value from the Census based on whether the households have heat, electricity and/or other services included in rent. Starting with reference year 2006 it will be possible to make this adjustment for owner households as well due to the addition of a new question in SLID asking whether such costs are included in other payments (such as condominium fees or property taxes); meanwhile, all owner households are imputed with positive utility costs.

Once the households in the Census are grouped with like records in SLID according to the input characteristics, three are chosen at random and the mean utility cost is then used as the value for the SLID respondent. If there are less than three households in the Census from which to take the mean value, the characteristics are collapsed to ensure that a single household's value is not used. For example, communities could be collapsed so that the seven nearest communities are combined.

Since the Census is only conducted every five years, the mean values of the Census year are updated annually using province-specific indexes of the utility prices paid by consumers.

Prior to the re-release of the housing data in SLID for 1999-2004 and the publication of this report, the mass imputation of utility costs in SLID was carried out using the annual Survey of Household Spending (SHS) as the source dataset. However, due to small sample sizes leading to a great deal of removing and collapsing of input characteristics in

order to obtain mean values, the imputed values were too middle-of-the-road to give an acceptable distribution of utility costs. It was decided that the SHS data source should be dropped in favour of the more precise mean values that could be obtained from the Census; price indexes would serve as a way of updating these mean values between Censuses. As a result, utility costs for 2002-2004 have been revised using Census data.

9. Where to find more information

All information related to the survey can be found on the Statistics Canada web site (www.statcan.ca) by going to *Survey of Labour and Income Dynamics – A Survey Overview*. Click on "Our products and services", "free internet publications" and "Personal finance and Household Finance".

Also at that location is the *Survey of Labour and Income Dynamics Electronic Data Dictionary*, which contains detailed information on a variable-by-variable basis organized into subject themes. At the time of this report, the variables on housing are accessed by going to the themes "Personal characteristics", then "Family and household characteristics", then "Dwelling, housing".

The Income Research Paper Series publishes each year the full SLID questionnaires. Look for the most recent title that begins, "Survey of Labour and Income Dynamics (SLID): Entry Exit Component Interview Questionnaire".

Additional information can be requested by contacting: Client Services, Income Statistics Division, 1-888-297-7355, fax: (613) 951-3012, e-mail: income@statcan.ca.

References

Canada Mortgage and Housing Corporation. 2000. "Special Studies on 1996 Census Data: Canadian Housing Conditions." *Research Highlights: Socioeconomic Series*, Issue 55-1.

Chris Li, Gaétan Garneau and Heather Lathe. 2005. *Recent Changes in Geography Content in the Survey of Labour and Income Dynamics (SLID)*. Income Research Paper Series. Catalogue no. 75F0002MIE2005002. Ottawa: Statistics Canada.

Duddek, Chris. 2006. "Modeling SLID Utility Costs using 2001 Census Data". May 24, 2006. Internal manuscript. Ottawa: Statistics Canada.

Duddek, Chris, and Susan Storey. 2006. *Data Quality in the 2004 Survey of Labour and Income Dynamics (SLID)*. Income Research Paper Series. Catalogue no. 75F0002MIE2006005. Ottawa: Statistics Canada.

House, Georgina. 2005. *General Housing Imputation (excluding utilities) in the Survey of Labour and Income Dynamics (SLID)*. Income Research Paper Series. Catalogue number 75F002MIE2005010. Ottawa: Statistics Canada.

Income Statistics Division. 2006. *Survey of Labour and Income Dynamics: A Survey Overview*. Income Research Paper Series. Catalogue no. 75F0011XIE2006001. Ottawa: Statistics Canada.

Storey, Susan. 2006. "Historic Housing Imputation Revision for the Survey of Labour and Income Dynamics Reference Year 1999-2004". April 24, 2006. Internal manuscript. Ottawa: Statistics Canada.